TCP-Planet: A New Reliable Transport Protocol for InterPlaNetary Internet

Ian F. Akyildiz, Ozgur B. Akan, Jian Fang

Broadband & Wireless Networking Laboratory
School of Electrical and Computer Engineering
Georgia Institute of Technology
Tel: 404-894-5141; Fax: 404-894-7883
Email: {ian, akan, jfang}@ece.gatech.edu
Web: http://www.ece.gatech.edu/research/labs/bwn

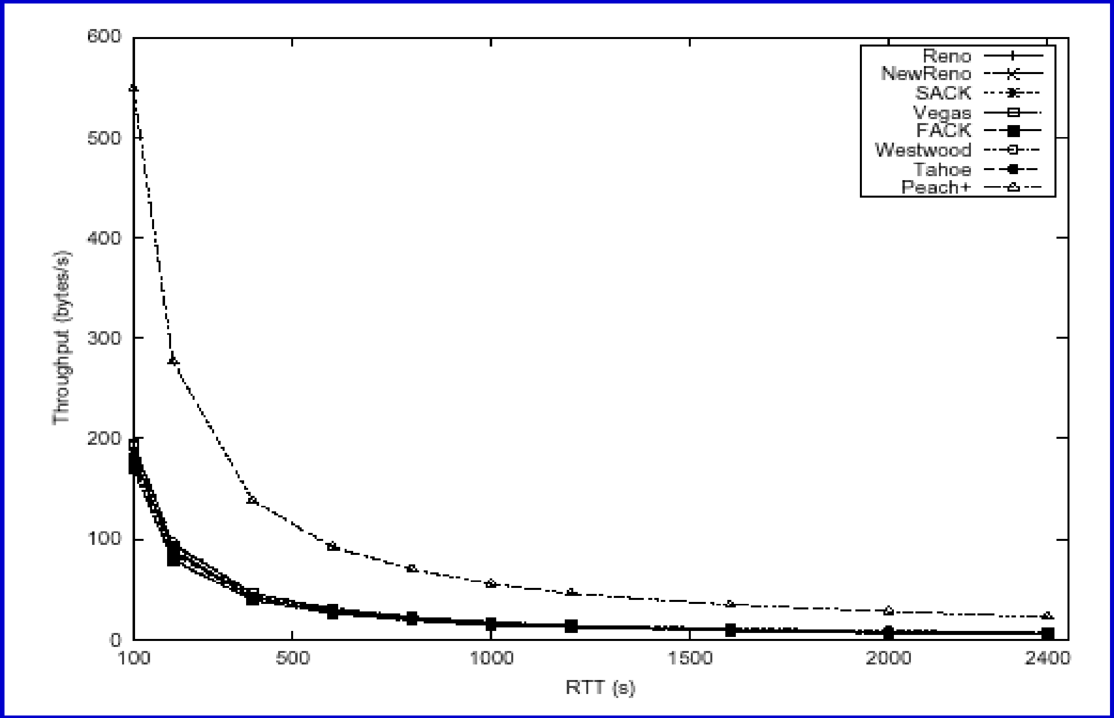
Challenges for Transport Layer

- Extremely High Propagation Delays
- High Link Error Rates
- Asymmetrical Bandwidth
- Blackouts

Performance of Existing TCP Protocols

- ***Window-Based*** **TCP is not suitable!!!**

  For *RTT = 40 min → 20B/s* throughput on *1Mb/s* link !!



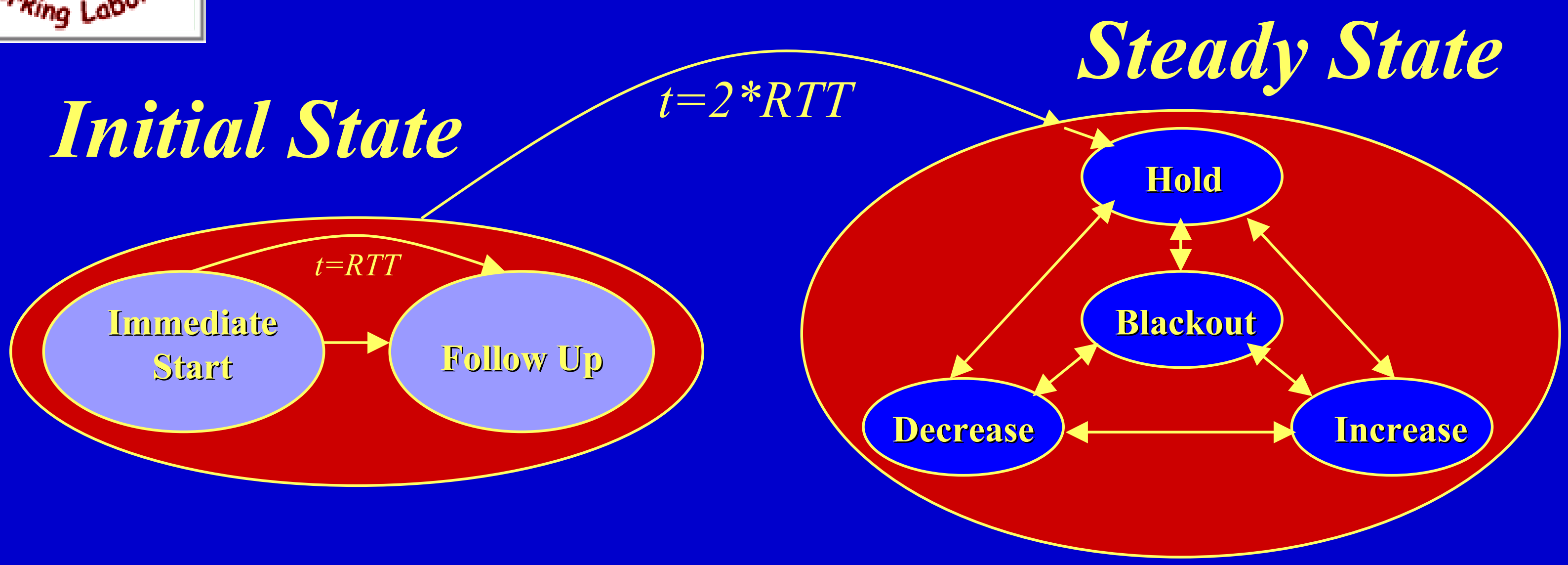
O. B. Akan, J. Fang, I. F. Akyildiz, "Performance of TCP Protocols in Deep Space Communication Networks",
***IEEE Communications Letters*, Vol. 6, No. 11, pp. 478-480, November 2002.**

Space Communications Protocol Standards – Transport Protocol (SCPS-TP)

- Addresses link errors, asymmetry, and outages
- SCPS-TP: Combination of existing TCP protocols:
  - Window-based
  - Slow Start
  - Retransmission Timeout
  - TCP-Vegas Congestion Control Scheme – variation of the RTT value as an indication of congestion!
- SCPS-TP Rate-Based:
  - Does not perform congestion control
  - Uses fixed transmission rate

New Transport Protocols are needed !!!

TCP-Planet



- ***Objective:*** **To address challenges of InterPlaNetary Internet**
- **A New *Initial State* Algorithm**
- **A New Congestion Detection Algorithm in *Steady State***
- **A New *Rate-Based* scheme instead of W*indow-Based***

***I. F. Akyildiz, O. B. Akan, J. Fang, "TCP-Planet: A Reliable Transport Protocol for InterPlaNetary Internet", *to appear in IEEE Journal of Selected Areas in Communications (JSAC),* early 2004.**

Initial State

- **Objective:**
  - * Overcome the Slow Start and RTT
  - * Determine → available bandwidth
  - * SET the initial transmission rate, S_{init}, ASAP !!!

- ***Immediate Start*** ***(Within one RTT)*** **:**
  - * Emulated Slow Start Phase
  - * Emulated Congestion Avoidance Phase

- ***Follow-Up*** ***(After (RTT + T) until 2 * RTT)***
  - Update transmission rate, S, according the feedbacks from the receiver

Initial State
(Immediate Start)

Two Basic Ideas:

- Divide *RTT* into small time intervals of length *T*

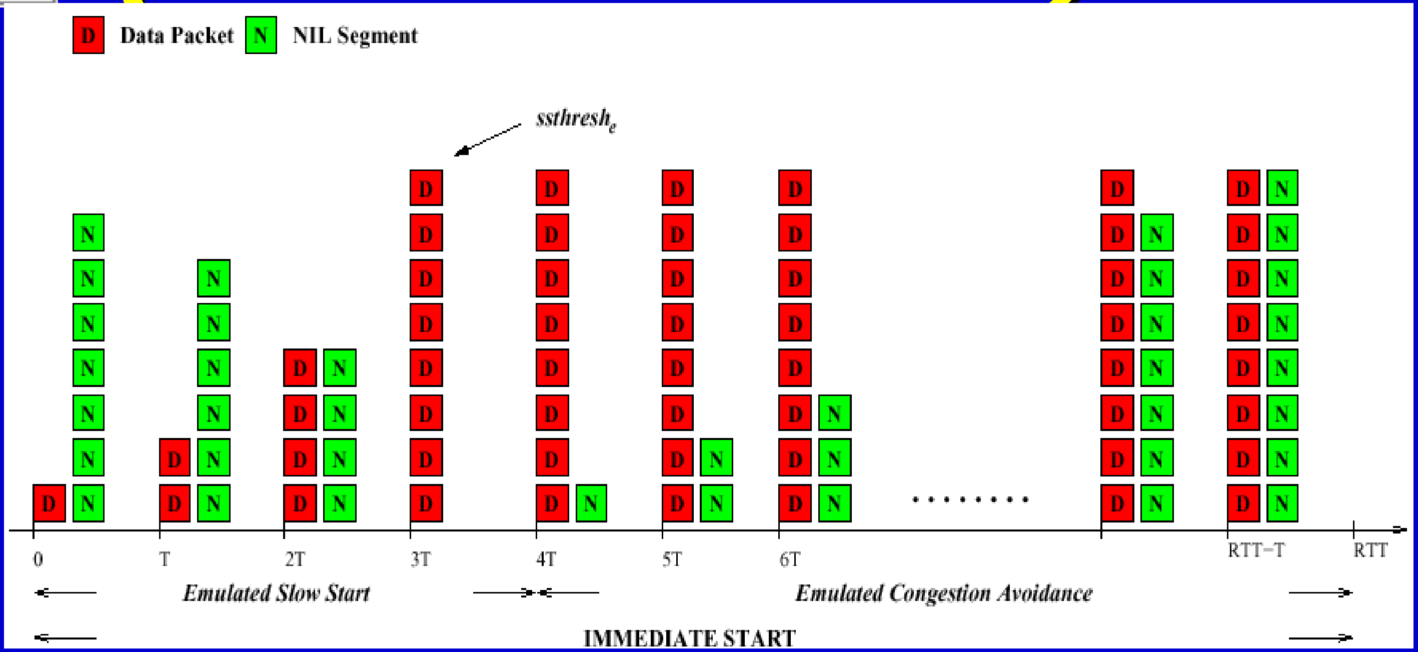
$$T = (RTT/B)^{1/2}$$

where *B* is the target transmission rate.

→ Use *T* as the RTT of the *emulated connection.*

- *Use NIL Segments (low priority) to probe network conditions.*

Initial State (Immediate Start)



- *Emulated Slow Start* :
  - Send total *$ssthresh_e$* packets in each *T* where *$ssthresh_e=(RTT*B)^{1/2}$*
  - Increase number of *data packets* and decrease number of NIL packets in *T* until *$ssthresh_e$* is reached.
- *Emulated Congestion Avoidance* :
  - Increase number of NIL packets in *T* until *t=RTT*.

Initial State *(Follow-Up)*

- Receiver counts and sends back the number of packets (N) it received in every T to the sender (in ACK packets).
- At $t=RTT+T$
  - Set Initial Data Rate $S_{init}=N/T$
- At each $t=RTT+k*T$ for $t<2RTT$ (feedback for k^{th} interval)
  - Update data rate $S=N_k/T$
- At $t=2RTT$
  - Move to Steady State (Hold State)

Steady State

- New *Congestion Control* Method to address link problems (also distinguish between congestion and link errors!!)
- New *Adaptive Rate-Based* AIMD protocol
- *Blackout State* to address the link outages
- *Delayed-SACK* to address the bandwidth asymmetry

Steady State
(Congestion Control)

- Source: Sends periodically *NIX segments*
  - Low and high priority
  - Small size, e.g., 40 bytes.
  - Both are sent with the same rate (= data rate).

Objective → Congestion decision based on the statistics of these received segments!

Fact 1. Since both are equal in size and both are sent with same rate, they may experience the same loss rate due to SPACE LINK ERRORS.

Fact 2: In case of congestion, low priority NIX segments are discarded → DIFFERENT LOSS RATES.

Steady State
(Congestion Control)

Let

$$\Phi := N_{low} / N_{high}$$

If $\Phi < 1$ → **CONGESTION!!!!**

NOTE:

Receiver does not ACK NIX segments

→ **but periodically sends back the number of low and high *NIX packets received, i.e.,* N_{low} and N_{high} *within a sliding time window.***

Steady State
(Congestion Control)

- Compare Φ with ϕ_d and ϕ_i
  (Congestion Decision Thresholds)

- If $(\Phi > \phi_i)$
  - Move to *Increase State* → $S=S+\alpha$

- If $(\Phi < \phi_d)$
  - Move to *Decrease State* → $S=S*\xi$

- If $(\phi_d < \Phi < \phi_i)$
  - Move to *Hold State* → $S=S$

Steady State
(Adaptive Rate-Based AIMD)

- **Default AIMD parameters *(α=1, ξ=0.5)* not suitable for long delay space links**
- **Throughput, T, depends on *RTT, p*, α, and ξ.**

$$T = \frac{\alpha}{4\cdot(1-\xi)}\left[1+\xi+\sqrt{(3-\xi)^2+\frac{8\cdot(1-\xi^2)}{\alpha\cdot RTT\cdot p}}\right]$$

- **New Adaptive Rate-Based AIMD scheme**
  - **Set ξ high from (1> ξ >0.5).**
  - **Adjust AIMD increase parameter α based on link conditions *RTT, p* to reach target throughput *B***

$$\alpha = \frac{(1+\xi)}{2}\left(B+\frac{1}{RTT\cdot p}\right)\left[\sqrt{1+\frac{8B^2(1-\xi)}{\left(B+\frac{1}{RTT\cdot p}\right)^2(1+\xi)^2}}-1\right]$$

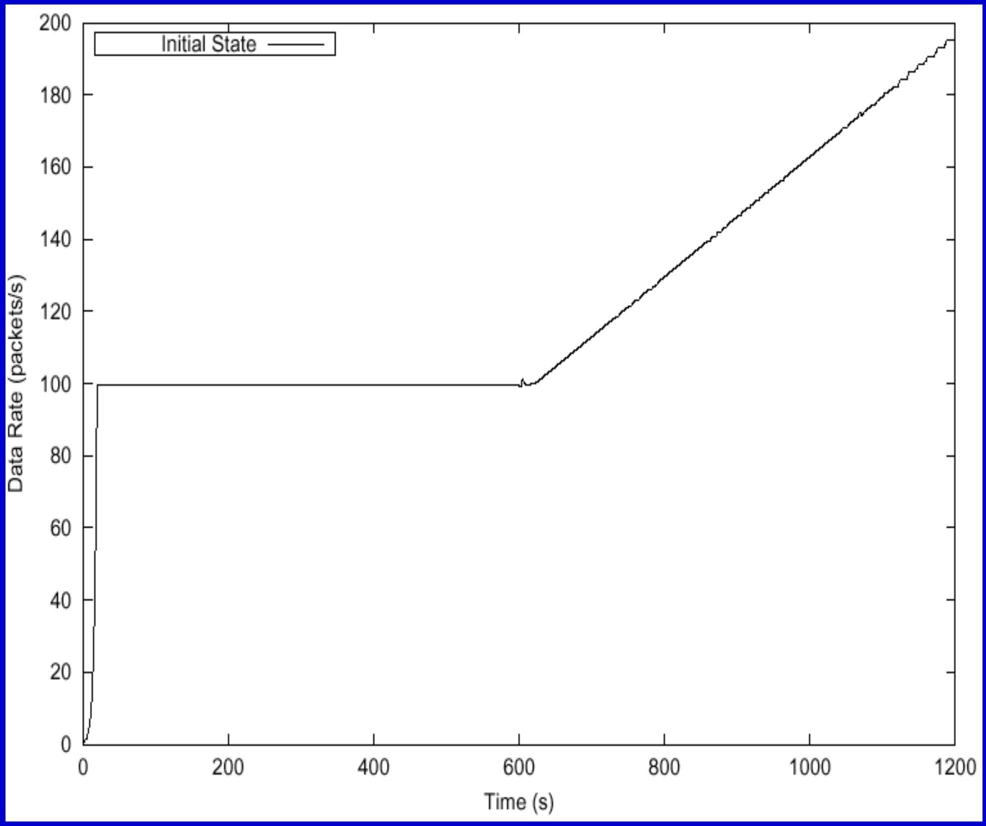
Steady State
(Blackout State)

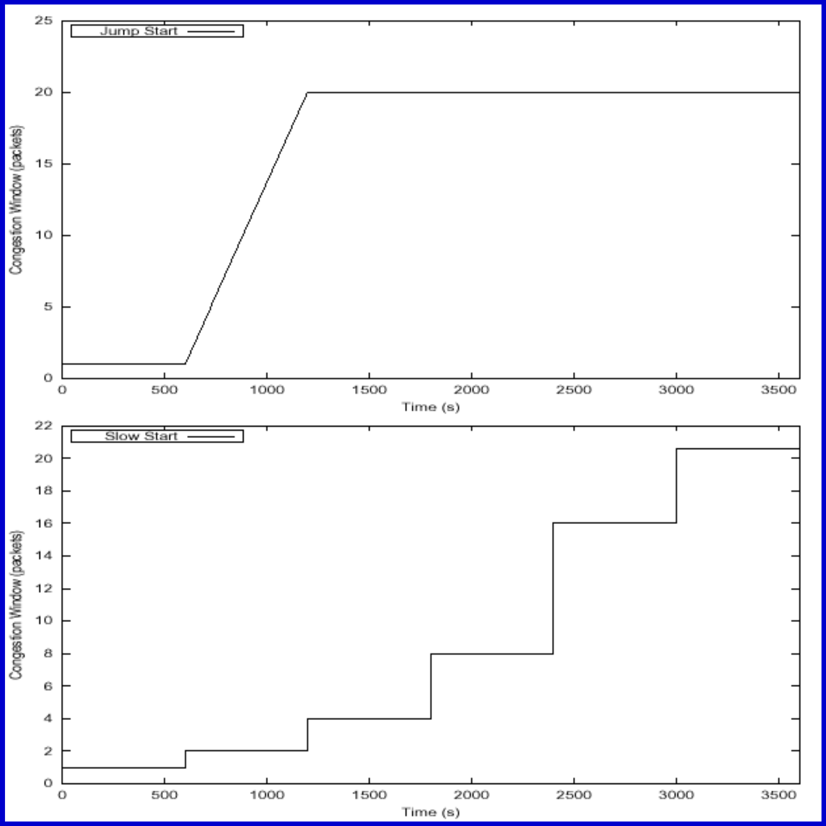
- If sender does not receive ACK for a certain period → goes to *Blackout State*
- Blackout State
  - *Objective:* To reduce the throughput degradation due to blackout.
  - Source stops sending new data packets but keeps sending NIX packets.
  - Receiver keeps sending NIX ACKs including N_{low} and N_{high} .
  - If the source starts to receive NIX ACKs with ($N_{low}=0, N_{high}=0$) → goes to *HOLD State.*
  - Otherwise, the new state is decided based on Φ .

Steady State *(Bandwidth Asymmetry)*

- ACK for each packet leads to congestion in the reverse link for asymmetrical space links
- Delayed-SACK
  - SACK is used for reliability.
  - Receiver delays SACK packets to avoid congestion in the reverse link.
  - If no packet loss, then receiver sends ***1 SACK*** per ***d*** (delayed-SACK factor) received packets.
  - If packet loss occurs (SACK fields need to be updated), receiver sends a new ***SACK*** immediately.

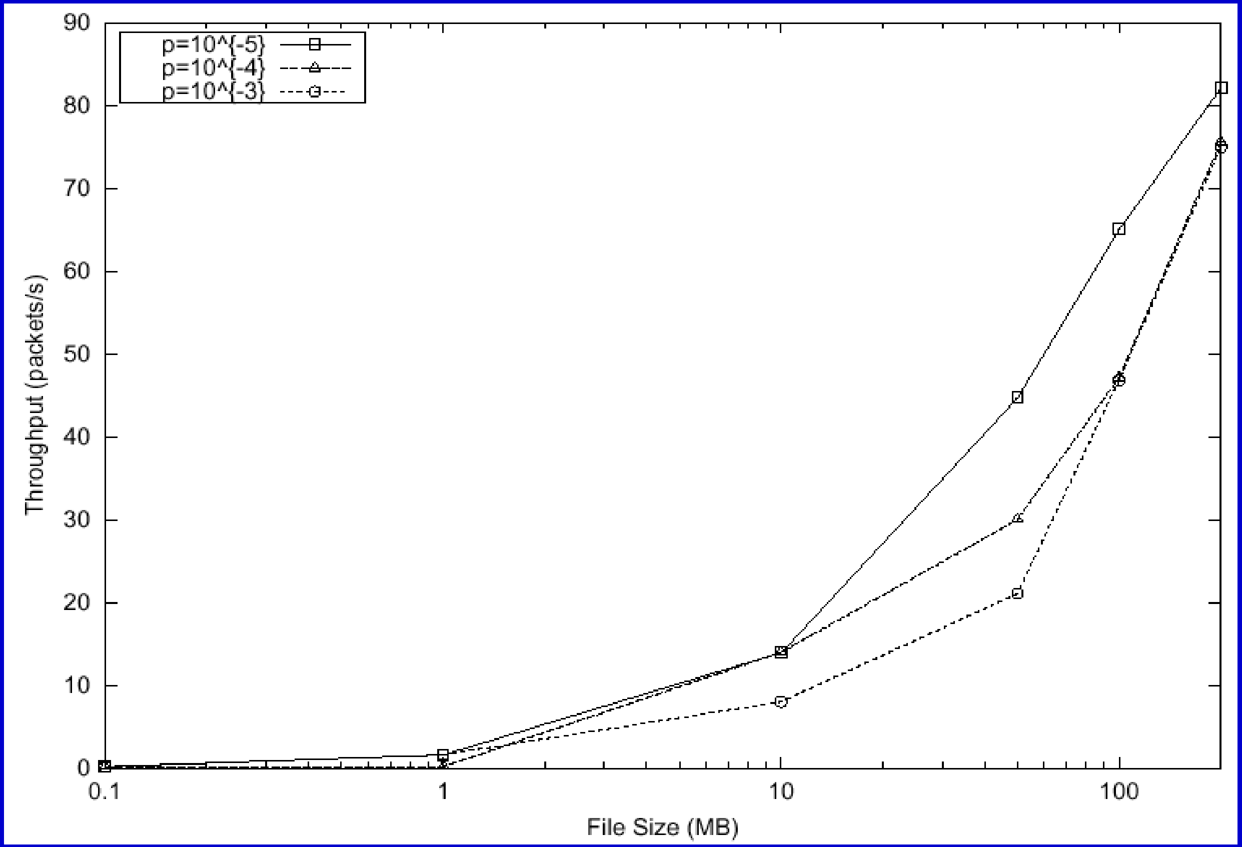
Performance Evaluation *(Initial State)*





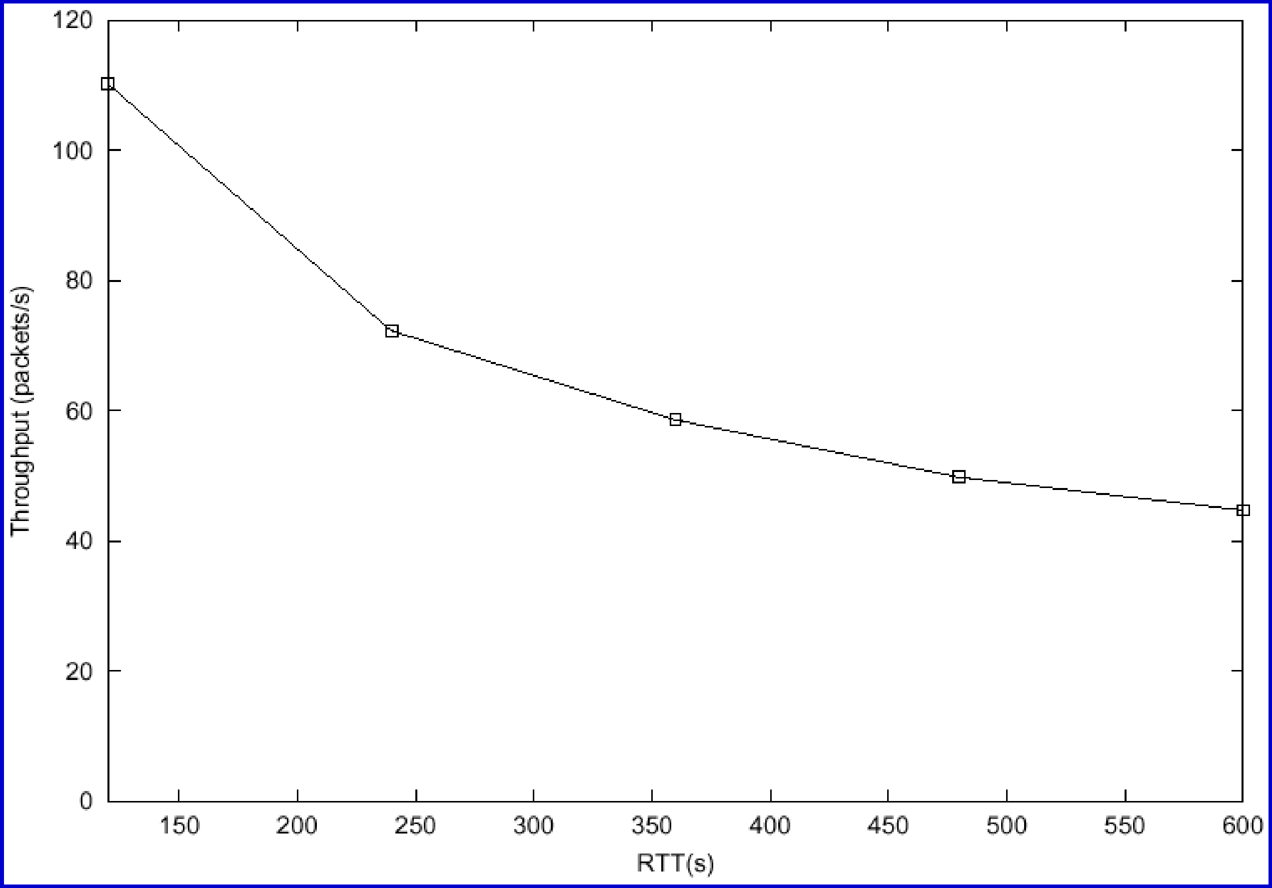
- Initial State (TCP-Planet) vs. Jump Start (TCP-Peach+) and Slow Start (TCP); RTT=600 sec; p=10^{-5}; Target Rate =100packets/sec.

Performance Evaluation *(Throughput)*



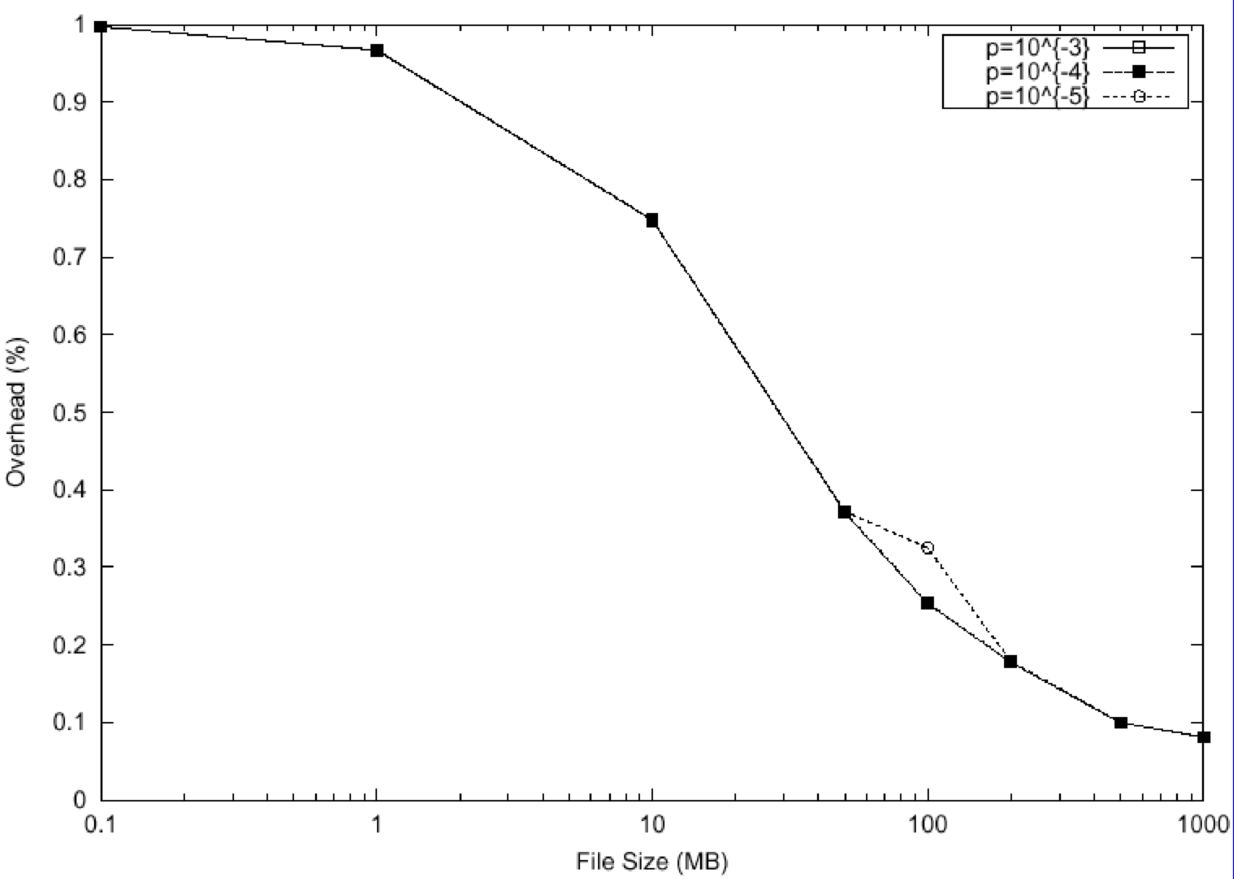
- **Throughput vs. File size;** ***RTT=600 s, p=10^{-5}, 10^{-4}, 10^{-3}***, **Link 1Mb/s; Target rate =** ***100 packets/sec (→ 100 KB/sec for data packets of size 1KB). NOTE: 200 MB → Vegas (SCPS-TP) → 30 B/sec;***
  → Planet → 83 KB/sec !!!!!

Performance Evaluation *(Throughput)*



- **Throughput vs. RTT; File Size *=50MB, p=10^{-5}, 10^{-4},10^{-3}*, Link Rate: 1Mb/s, Target Rate = *100 KB/s***

Performance Evaluation *(Overhead)*



- **TCP-Planet Overhead due to NIL (in Initial State) and NIX (in Steady State) packet transmissions. Varying file size, RTT=600 s, *$p=10^{-5}$, 10^{-4}, 10^{-3}*, link 1Mb/s, Target rate = *100 KB/s***

Conclusions

- Existing TCP protocols not suitable for InterPlaNetary Internet
- TCP-Planet for InterPlaNetary Internet
  - New Initial State instead of Slow Start
  - New Rate-based AIMD instead of window-based
  - New Congestion Control to address link errors
  - Blackout State to address link outages
  - Delayed SACK for bandwidth asymmetry

Performance evaluation shows TCP-Planet addresses the challenges and significantly improves the network performance.

I. F. Akyildiz, O. B. Akan, J. Fang, "TCP-Planet: A Reliable Transport Protocol for InterPlaNetary Internet", *to appear in IEEE Journal of Selected Areas in Communications (JSAC),* early 2004.